CMB Power Spectrum Estimation with Hamiltonian Sampling

J. F. Taylor, M. A. J. Ashdown, M. P. Hobson
Astrophysics Group. Cavendish Laboratory. JJ Thomson Avenue. Cambridge CB3 0HE, UK

We present a method for fast optimal estimation of the temperature angular power spectrum from observations of the cosmic microwave background. We employ a Hamiltonian Monte Carlo (HMC) sampler to obtain samples from the posterior probability distribution of all the power spectrum coefficients given a set of observations. We demonstrate the method on simulated WMAP observations. We then discuss extending the method to handle polarisation and demonstrate how sampling avoids the problem of E/B mixing caused by incomplete sky coverage.

1 Introduction

Observations of the cosmic microwave background (CMB) have proved to be extremely valuable for testing and constraining cosmological models. The majority of models predict that the anisotropies in the CMB signal are Gaussian and their statistics isotropic across the sky. The angular power spectrum C_ℓ therefore provides a natural connection between theory and observation and a variety of methods have been explored to compute the power spectrum from sets of observations.

Maximum-likelihood methods[5,1,9] provide an optimal estimate of the CMB power spectrum however brute force implementations of the method can only be applied to small data sets as the required computation scales as $\mathcal{O}(N_{\rm pix}^3)$, where $N_{\rm pix}$ is the number of pixels in a CMB map. Alternatively one can resort to approximate pseudo-C_ℓ methods[7]. These scale as the map-making process and are fast even for the largest data sets. Both types of method can only supply an approximation to the likelihood function required to compare spectra predicted from theory with those estimated from observations.

An alternative framework has been developed[12,8] where one explores the full posterior distribution of the power spectrum with Monte Carlo samples. This method is not only exact but scales like the pseudo-C_ℓ methods. Under the assumption of position invariant, circularly symmetric beams and uncorrelated noise, the method scales as $\mathcal{O}(N_{\rm pix}^{3/2})$.

The approach relies on the availability of an efficient method for sampling from high-dimensional distributions. Previous implementations use a Gibbs sampler but this restricts the applicability of the method to Gaussian noise and CMB. We propose the use of a Hamiltonian Monte Carlo (HMC) sampler[3]. As opposed to the majority of Markov-Chain Monte Carlo (MCMC) methods, HMC scales well with problem size. Few requirements are made on the distribution to be sampled, thus giving us the opportunity for great flexibility.

2 Power spectrum estimation with sampling

Suppose the true CMB sky, described by a pixelised map, is represented by the vector $\mathbf{t}$. The sky is observed and the resultant data vector (in any domain) $\mathbf{d}$ is the sum $\mathbf{d} = \mathbf{s}+\mathbf{n}$ of contributions due to the underlying CMB $\mathbf{s}$ and noise $\mathbf{n}$. Moreover the signal is usually related to the true sky by the linear mapping $\mathbf{R}$. The field $\mathbf{t}$ is related to the spherical harmonics by

$$t(x_p) = \sum_{\ell=2}^{\ell_{\max}} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(x_p), \tag{1}$$

where $t(x_p)$ is a single pixel in the map and $Y_{\ell m}$ are the spherical harmonics. In this notation we may write our model for the data as

$$\mathrm{d} = \mathbf{RY}\mathrm{a} + \mathrm{n}. \tag{2}$$

For an isotropic Gaussian CMB the covariance matrix $\mathbf{C}$ of the $a_{\ell m}$ has components

$$C_{\ell m \ell' m'} = \langle a_{\ell m} a^*_{\ell' m'} \rangle = C_\ell \delta_{\ell\ell'} \delta_{mm'}, \tag{3}$$

where the set of $\{C_\ell\}$ coefficients constitute the angular power spectrum.

We aim to sample from the joint distribution $\Pr(\{C_\ell\}|\mathbf{d})$. Although this is difficult to perform directly, it is possible to sample from the joint density of the power spectrum coefficients and the signal realization $\Pr(\{C_\ell\}, \mathbf{a}|\mathbf{d})$ and then marginalise over $\mathbf{a}$. The joint density can be written as the product of the appropriate conditional distributions

$$\Pr(\{C_\ell\}, \mathbf{a}|\mathbf{d}) \propto \Pr(\mathbf{d}|\mathbf{a}) \Pr(\mathbf{a}|\{C_\ell\}) \Pr(\{C_\ell\}). \tag{4}$$

Assuming a flat prior ($\Pr(\{C_\ell\}) = 1$) and Gaussian noise then the conditional distributions that make up (4) can be written as

$$\Pr(\mathbf{d}|\mathbf{a}) \propto \exp\left[-\frac{1}{2}(\mathbf{d} - \mathbf{RYa})^{\mathrm{T}} \mathbf{N}^{-1} (\mathbf{d} - \mathbf{RYa})\right], \tag{5}$$

where $\mathbf{N} = \langle \mathbf{nn}^{\mathrm{T}} \rangle$, and

$$\Pr(\mathbf{a}|\{C_\ell\}) \propto \prod_\ell \left(\frac{1}{C_\ell}\right)^{\frac{2\ell+1}{2}} \exp\left(-\frac{2\ell+1}{2}\frac{\sigma_\ell}{C_\ell}\right), \tag{6}$$

where $\sigma_\ell = \frac{1}{2\ell+1}\sum_m |a_{\ell m}|^2$.

The selection of the domain in which to represent the data is determined by the requirement that $\mathbf{N}$ has a simple form. In this work we assume that in the map domain $\mathbf{N}$ is well represented by a diagonal matrix.

We draw samples from the joint space $(\mathbf{a}, \{C_\ell\})$ using a Hamiltonian Monte Carlo sampler.

3 Hamiltonian Monte Carlo

Hamiltonian Monte Carlo draws parallels between classical dynamics and sampling to suppress the random walk behaviour inherent in most MCMC techniques. To sample from the multidimensional distribution $\Pr(\mathbf{x})$ we introduce a momentum $\mathbf{p}$ and mass $\mathbf{m}$ and define the Hamiltonian

$$H = \frac{p^{\mathrm{T}} p}{2m} - \log \Pr(\mathbf{x}). \tag{7}$$

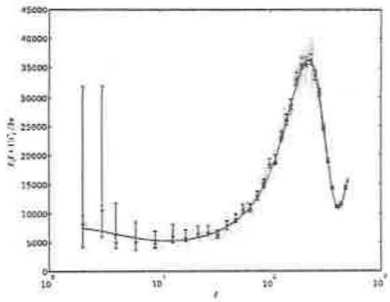

Figure 1: Binned power spectrum and 68 percent confidence intervals for the HMC sampling method (black circles and error bars) compared to results from the MASTER method (grey squares and errorbars) when applied to a simulated WMAP observation. the input spectrum for the simulation is shown in black and the spectrum of the realization in grey.

A new sample is proposed by drawing a set of random momenta $\mathbf{p}$ from a Gaussian with variance $\mathbf{m}$. We then move through the $(\mathbf{x}, \mathbf{p})$ space using Hamilton's equations (usually using a straightforward first order discretisation). After a random length of time $t = T$ we accept with probability p_A given by a modified Metropolis rule

$$p_A = \min\left[1, \exp(H(t = 0) - H(t = T))\right] \tag{8}$$

Since trajectories that obey Hamilton's equations conserve H we expect a high acceptance probability even after moving considerable distances in the parameter space.

4 Application to simulated WMAP observations

We produced a map of the CMB with a HEALPix[a] $N_{\rm side} = 512$ ($\sim 3 \times 10^6$ pixels). Our CMB simulation is a realization of a ΛCDM cosmology with the best fitting parameters from the 5-year *WMAP* observations[b] [10] and includes multipoles up to $\ell = 512$.The map was then smoothed with a 13-arcmin Gaussian beam, which is similar in size to the beam of the *WMAP* *W*-band. We then added anisotropic uncorrelated noise by making use of the published[c] $N_{\rm obs}$ and noise variance for the 5-year *WMAP* combined *W* band map. The map was cut with the Kp2 mask which excludes 15.3% of the sky. We included multipoles up to $\ell_{\rm max} = 512$ in our analysis. This gives us a total of around 2×10^5 parameters in our sampling space.

For these simulations we made a total of 5000 burn in samples and recorded 10000 samples from the post burn-in phase. It takes around 20 seconds to generate a single sample using two dual core Intel Xeon 5150 processors and the MPI parallelised HEALPix spherical harmonic transforms, resulting in a total processing time of around 80 hours.

For comparison we applied the MASTER method [7] to the same data set. Our peak likelihood C_ℓ sample and 68 per cent confidence intervals, binned with the *WMAP* team's scheme, are shown alongside the results of the MASTER method in Fig.1. For most of the range of angular scales the two estimates and their errors agree well. On the largest angular scales the MASTER estimate tends to underestimate the uncertainties and the symmetric errors are far from representative of the posterior.

[a] http://healpix.jpl.nasa.gov

[b] http://lambda.gsfc.nasa.gov/product/map/dr3/parameters.cfm

[c] http://lambda.gsfc.nasa.gov

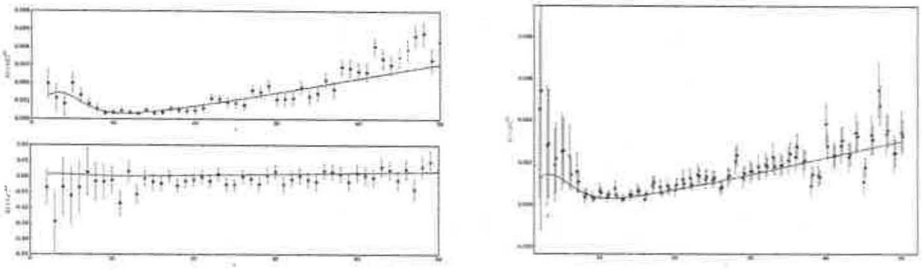

Figure 2: The plots on the left show the BB spectrum estimated using Spice with the full sky (top) and when the data is cut with the WMAP polarisation mask (bottom). The plot on the right shows the results of the sampler applied to the same simulations, fullsky in blue and cut sky in red.

5 Extending the method to polarisation

Polarisation provides a new set of opportunities for science and challenges for data analysis. Of particular interest is a measurement of the primordial B mode spectrum predicted by many inflationary models. One problem faced in the analysis is that for observations on the cut sky the separation into E and B modes is ambiguous. The E/B mixing is a problem for pseudo-C_ℓ methods whereas exact methods, such as sampling, can still construct optimal estimates. We apply Spice [11,2] and the HMC method to a set of signal dominated simulations. Figure 2 shows how Spice fails to recover the BB when a cut is imposed whereas the sampler is relatively unaffected, a slight increase in the uncertainties caused by the removal of data from the cut.

Acknowledgements

We thank Morgan French for his contributions to the early development of our sampler. JFT acknowledges a STFC (formerly PPARC) studentship. MAJA is a member of the Cambridge Planck Analysis Centre, supported by STFC grant ST/F005245/1 .This work was conducted in cooperation with SGI/Intel utilising the Altix 3700 supercomputer at DAMTP Cambridge supported by HEFCE and STFC. We acknowledge the use of the Legacy Archive for Microwave Background Data Analysis (LAMBDA). Support for LAMBDA is provided by the NASA Office of Space Science. Some of the results in this paper have been derived using the HEALPix [6] package.

1. Bond J., Jaffe A. H., Knox L., 1998, Physical Review D, 57, 2117
2. Chon et al, 2004, MNRAS 350, 914
3. Duane S., Kennedy A., Pendleton B. J., Roweth D., 1987, Physics Letters B, 195, 216
4. Efstathiou G., 2003, MNRAS, 346, L26
5. Gorski K. M., 1994, ApJ, 430, L85
6. Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, ApJ, 622, 759
7. Hivon E., Górski K. M., Netterfield C. B., Crill B. P., Prunet S., Hansen F., 2002, ApJ, 567, 2
8. Jewell J., Levin S., Anderson C., 2004, ApJ, 609, 1
9. Oh S. P., Spergel D. N., Hinshaw G., 1999, ApJ, 510, 551
10. Spergel D. N. et al., 2007, ApJS, 170, 377
11. Szapudi, Prunet, Colombi, 2001, ApJ 548, 115
12. Wandelt B. D., Larson D. L., Laksminarayana A., 2004, Physical Review D, 70, 083511